

This document is an English translation of the abstract of Korean Patent Application Publication No. 1997-0056108.

Abstract

The present invention relates to a method of compressing Korean alphabet. It has advantages in that compressibility is improved by re-organizing dictionary tree inner table to recognize the code, after converting array of an initial consonant, vowel and a final consonant into code number by rearranging the array according to probability of generation, and transmission rate is improved by compressing Korean alphabet before transmitting Korean alphabet data packet.

(19) 대한민국특허청(KR) (12) 공개특허공보(A)

(51) Int. Cl. ⁶ H04J 3/18	(11) 공개번호 (43) 공개일자	특 1997-0056108 1997년 07월 31일
(21) 출원번호 (22) 출원일자	특 1995-0067428 1995년 12월 29일	
(71) 출원인	사단법인 고등기술연구원 연구조합 이우복	
(72) 발명자	서울특별시 중구 남대문로 5가 541 대우센터 빌딩 (우:100-095) 서두원	
(74) 대리인	경기도 군포시 산본동 백두한양아파트 989동 501호 손경한	

심사청구 : 있음

(54) 조합형 한글의 압축방법

요약

본 발명은 조합형 한글의 압축방법으로서, 초성, 중성, 종성의 배열을 발생활자의 빈도에 따라 재배열시킨 코드번호로 변환시킨 후 상기 코드를 인식할 수 있도록 딕셔너리 트리 내부 테이블을 재구성하여 항상된 압축률이 좋아지며 한글 데이터 패킷의 전송에 앞서 한글 압축 단계를 거치게 함으로서 전송률을 높일수 있다.

대표도

도 4

명세서

[발명의 명칭]

조합형 한글의 압축방법

[도면의 간단한 설명]

제4도는 조합형 한글의 내부 테이블의 구조도,

제5도는 조합형 한글의 플로우차트도.

본 내용은 요부공개 건이므로 전문 내용을 수록하지 않았음

(57) 청구의 범위

청구항 1

한글의 초성, 중성, 종성의 코드표를 재구성하는 단계와, 상기 재구성된 코드표를 아스키코드를 위한 128개의 노드와 초성, 중성, 종성의 자소코드번호 32개의 노드를 딕셔너리 트리내 초기화하는 단계와, 상기 초기화된 압력문자를 읽어 딕셔너리 트리내에 있는 스트링으로 취하는 단계와, 상기 스트링 단계후 다음 입력문자가 있는가를 판단하여 입력문자가 있으면 다음문자를 읽어 싱글 캐릭터로 하고 입력문자가 없으면 딕셔너리 트리내에 있는 스트링에 할당한 부호어를 출력한 후 종료하는 단계와, 상기 딕셔너리 트리내에 있는 스트링과 싱글 캐릭터가 딕셔너리 트리내 존재하는가를 판단하여 딕셔너리에 존재하면 딕셔너리 트리내에 있는 스트링과 싱글 캐릭터를 새로운 스트링으로 하고 딕셔너리에 존재하지 않으면 딕셔너리 트리내에 있는 스트링에 할당한 부호어를 출력하는 단계와, 상기 출력된 딕셔너리 트리내에 대한 새로운 노드를 생성하여 부호어를 할당하여 다음 싱글 캐릭터를 새로운 스트링으로 취하는 단계를 포함하는 조합형 한글의 압축 방법.

청구항 2

제1항에 있어서, 상기 조합형 한글코드는 1비트의 MSB와 5비트의 초성 중성, 종성으로 구성되어 있는 것을 특징으로 하는 조합형 한글의 압축방법.

청구항 3

제2항에 있어서, 상기 MSB는 한글인지 아스키 코드인지를 구분하는 것을 특징으로 하는 조합형 한글의 압축방법.

※ 참고사항 : 최초출원 내용에 의하여 공개하는 것임.

도면

도면4

0 - 127	ASCII CODE
128 - 159	한글 코드
160 - 255	1비트 부호화 영역
256 - 511	2비트 부호화 영역
512 - 1023	3비트 부호화 영역
1024 - 2047	4비트 부호화 영역
2048 - 4095	5비트 부호화 영역
4096 - 8191	6비트 부호화 영역

도면5

